DOCUMENT RESUME

ED 482 495 JC 030 651

AUTHOR	Barr, James E.; Rasor, Richard A.; Grill, Cathie
TITLE	The Evaluation of Present Course Placement Procedures Using the Compass Tests.
INSTITUTION	American River Coll., Sacramento, CA.
PUB DATE	2002-01-00
NOTE	44p.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	EDRS Price MF01/PC02 Plus Postage.
DESCRIPTORS	Academic Standards; Cognitive Processes; Community Colleges; Comparative Testing; Content Analysis; Evaluation; Objectives; Skill Analysis; *Task Analysis; Test Construction; *Two Year Colleges
IDENTIFIERS	*American River College CA

ABSTRACT

This document addresses how well ARC's computerized placement tests (Compass) assist individuals in reaching informed decisions about enrolling in selected courses, including English composition, reading, mathematics, and ESL. The document addresses the question of whether Compass scores add any relevant information in the decision-making process surrounding course placement. To answer this question, evidence is considered based upon the test content, relations to future external variables, and consequences of testing, convergent data, and professional judgment. Data is analyzed for each subject test and tables of correlational data is provided for coefficients between Compass scores and grades under varying data conditions for English, Math, Reading and correlation coefficients between Compass scores and grades for ESL. The study concludes that the original correlations are either very low, non-existent, or even negative. This means that counselors who believe that Compass tests should be given heavy weight are misguided and that entry decisions based partly upon Compass test outcomes are of dubious values. A possible solution is to have instructors of the same course write specific entrance skills that all students must have as a prerequisite to enrolling in a course. The document concludes with a list of implications for planning. (MZ)

Reproductions supplied by EDRS are the best that can be made
from the original document.

C:\researchweb6\Documents\Student Self Placement.doc
Created on 2/4/2002 9:50 AM

The Evaluation of Present Course Placement Procedures Using the Compass Tests

American River College

ED 482 495

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

James E. Barr
Richard A. Rasor
Cathie Grill

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

R. A. Rasor

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

January 2002

IC030651

BEST COPY AVAILABLE

Table of Contents

INTRODUCTION TO ARC COMPASS ASSESSMENT TEST PERFORMANCE **1**
English Placement Test (Compass) for English 1A, 58, 256 ..... ***4***
English Writing Cut Scores 4
ARC Normative Data 4
English 1A Analysis 5
English 58 Analysis 6
English 256 Analysis 7
Reading Placement Test (Compass) for English 4, 5, 71 (now 78), & 268/270 ..... ***8***
English Reading Cut Scores 8
ARC Normative Data 8
English 4/5 Analysis 9
English 71 Analysis (Course number has now been changed to 78) 10
English 268/270 Analysis 11
Math Placement Test (Compass) for Math 51, 53, 15 ..... ***12***
Math Cut Scores 12
ARC Normative Data for Math Placement (Compass) 12
Math 51 Analysis 15
Math 53 Analysis 16
Math 15 Analysis 17
ESL Placement Test (Compass) for ESL 260, 270, 280, 1, 2, 63, 5 ..... ***18***
ESL Cut Scores 18
ARC Normative Data 18
Eight Levels of ESL Course Placement with Percentiles 20
Recently Implemented: Seven Levels of ESL Course Placement with Percentiles 20
ESL 2 R (Reading) Analysis 21
ESL 1 R (Reading) Analysis 22
ESL 280 R (Reading) Analysis 23
ESL 270 R (Reading) Analysis 24
ESL 260 R (Reading) Analysis 25
ESL 2 W (Writing) Analysis 26
ESL 1 W (Writing) Analysis 27
ESL 280 W (Writing) Analysis 28
ESL 270 W (Writing) Analysis 29
ESL 260 W (Writing) Analysis 30
ESL 2 L (Listening) Analysis 31
ESL 1 L (Listening) Analysis 32
ESL 280 L (Listening) Analysis 33
ESL 270 L (Listening) Analysis 34
ESL 260 L (Listening) Analysis 35
Tables of Correlational Data ***36***
Table 1. Correlation Coefficients Between Compass Scores and Grades Under Varying Data Conditions For English, Math, Reading 37
Table 2. Correlation Coefficients Between Compass Scores and Grades For ESL 38
SUMMARY OF THE COMPASS ASSESSMENT TEST OUTCOMES **39**
IMPLICATIONS FOR PLANNING **40**

List of Figures

Figure 1. All Compass English Writing Assessment Scores 4
Figure 2. All Compass Reading Assessment Scores 8
Figure 3. ARC Normative Distributions for Five Compass Math Tests. 14
Figure 4. ARC Normative Distributions for Three Compass ESL Tests. 19

INTRODUCTION TO ARC COMPASS ASSESSMENT TEST PERFORMANCE

This report deals with one aspect of institutional effectiveness, how well ARC's computerized placement tests (Compass) assist individuals in reaching informed decisions about enrolling in selected courses, namely English composition, reading, mathematics, and ESL. The bottom line question is: Do Compass scores add any relevant information in the decision-making process surrounding course placement? If so, to what degree? The research to answer these questions is a matter of validity. Of critical concern is the interpretation of evidence purporting to show the degree of validity. Validity in this case is the defensibility of the inferences made on the basis of test scores and other measures with respect to student performance in English, reading, mathematics, and ESL. In other words, it is the accuracy of total judgments leading to a decision of course entry.

The wording "test validity," while commonly used throughout educational circles, is really an inappropriate term. Rather, ***it is the decisions made using test outcomes that must be validated.*** There are several such sources of validity evidence. The main sources used in educational settings such as ours are briefly described below:

1) Evidence based on test content. Often this is the judgment of professionals in the field who can examine in detail the congruency between the test items and the construct it is intended to measure, e.g., content of a course. Thus a test score would enable one to determine the amount of knowledge that a student has about the subject matter before taking such a course. Given this objective, it would seem logical that scores only be used to skip a target course and proceed to the next level. Students who start a course of study are not supposed to already know very much about the learning tasks awaiting them. If they do, then they should go to the next level course. What is necessary information is how well students perform on the entrance standards of a course - not upon the content of the course.

2) Evidence based on relations to future external variables. This evidence most often measures the degree of correlation between assessment test scores and future criterion measures like final grades in a course or scores or a comprehensive final exam within a course. The critical question is how well assessment scores predict another related external measure.

3) Evidence based on consequences of testing. This type of evidence usually speaks to the benefit gained by use of a testing system. For example, it would be beneficial if such testing were to result in higher student course performance through differential screening of individuals who lack the prerequisite skill to enter a course. In this instance, one must show evidence that such a benefit actually occurs. In the California community college system, there is often much discussion over what is called "Consequential Validity." This is often measured early in a semester by the percentage of students judging that they were placed correctly in a course with respect to the skills demanded of them to complete that course successfully. In addition, the instructor of the course also evaluates each student with respect to the same standard, that is, the student either was or was not properly placed. Although this type of validity evidence is used in several community colleges, and is thought by some to be a legally defensible position in terms of using tests, it seems to us that it misses the mark when it comes to what is intended by the consequences of testing. Consequences appropriate here would be improved student performance, improved teaching, increased student motivation, and decreased student dropouts - not simply that the instructor and the students in a course believe proper placement has occurred.

4) Evidence based on convergent data. Relationships between test scores and other measures intended to assess similar constructs provide convergent evidence. For example, scores on a multiple-choice test that measure correct grammatical usage of English may be compared with evaluations of correct usage within written essays. Experimental designs may also be used to determine if test scores change as a function of receiving instruction in a course. For example, a test that measures reading comprehension

may be given at the beginning of a course intended to affect reading skills, and again at the end of the course.

5) Evidence based on professional judgment. It is a legal requirement in California's community colleges that no decision to allow or disallow enrollment in a particular course of study be based upon a sole measure coming from a test. Rather, other measures must be applied (multiple measures) which can either support or not support the interpretation coming from a test score. The intent here is to look at several sources of data by which to make an informed decision about enrollment in a course. However, it is not enough to simply apply the multiple measures in reaching a consensus decision, but to also evaluate the evidence as to whether the other measures are relevant to making a correct decision. The totality of data must be examined for its accuracy. It is suggested that a portfolio of validity evidence be available for each test used in the context of course placement.

Listed in this report are the cutoff scores, normative data about each test, and evidence of validity for the tests used in conjunction with course placement in English composition, reading, mathematics, and English as a second language (ESL). It is our understanding that each test has previously been reviewed by faculty for content validity, that is, test items adequately sample knowledge of the same constructs found within the course of study. Yet there is little rationale why students should be expected to know the material covered in a course when they have not yet enrolled. (See comments under #1 above). We also assume that professional personnel in every case have applied multiple measures - even though those multiple measures have not been evaluated for their appropriateness or consistency of application.

The evidence of validity used within this report compares test scores with subsequent final grade outcomes (degrees of success) shown in a variety of data configurations. As such, the evidence shown is criterion related and predictive, i.e., predictive validity. We also briefly examined the outcomes should a particular cut score be raised a few points. None of the conclusions changed because of such a shift. In addition, we cannot determine just what set of multiple measures were applied nor whether they improved or made worse the prediction of course success from knowledge of test score. Thus the correlation values between scores and grades that will be reported are not based upon scores alone but upon scores plus some set of other measures that counselors used to help with the decision making process regarding course entry. Finally, we acknowledge that cut scores have recently been shifted upwards for ESL. However, statistical analysis of the new cut scores cannot be adequately done until students enter under the new entry rules.

Before starting the analysis of score data with statistical terms, we provide a brief recap of the statistical concepts that are most important in this context.

Measures of Central Tendency & Skew: The mean is the arithmetic average obtained by adding all scores and dividing by the number of cases. As such, it is like a measure of balance of a distribution of scores. Extremely low or high scores can pull the value of the mean up or down. If there is an abundance of very low scores without a similar number of high values, the distribution is said to be negatively skewed, the tail of the plotted distribution is on the left side. The exact opposite situation in which there is a tail of high scores is referred to as positive skewness. If any distribution has greater skewness than .500, the shape of the distribution is not close to being normal or bell-shaped. Another measure of centralness is the median, the midpoint or middle area in a distribution that has already been rank ordered from low to high. Medians are unaffected by extreme scores. Therefore, if a distribution has either high negative or high positive skewness, the median should be used in place of the mean as a measure of central tendency.

Percentile Points: A given percentile point or rank corresponding to a specific score indicates the percentage of the sample that had scores which fell at and below that point. For example, suppose that a

score of 62 is equal to the 75th percentile. Then 75% of the sample had a score equal to or below the score of 62. The 50th percentile cuts the ranked distribution in half and is therefore also the median. Percentiles merely transform scores into ranks within a 1 - 100 scale.

Correlation Coefficients & Statistical Significance: Correlation values range from –1.00 through .00 to +1.00. Negatives or near zero correlations between scores and grades are usually not what a researcher wants to find. In student assessment, researchers would like to see high positive correlation values between test scores and grades, a value of at least .40 which would indicate that higher scores tend to be associated with higher grades. Likewise, lower scores tend to be associated with lower grades. If a correlation value is very low or zero, the two measures are unrelated in any mathematical sense and one cannot predict the other.

When someone views a correlation value and alongside to the right sees one of the following ($p < .05$) ($p < .01$) ($p < .001$) or (ns), it refers to the probability that the obtained sample correlation came from a population of values with a true correlation value of zero (.00). Therefore, $p < .05$ means that one is about 95% confident (and 5% unsure) that the population correlation is not .00. When you see $p < .01$, it means that one is about 99% confident (and 1% unsure) that the population correlation is not .00. When you see $p < .001$, it means that one is about 99.9% sure that the population correlation is not .00. Finally, when you see (ns), it indicates that one believes the obtained correlation value is likely to be nothing but random variation from a population with a true correlation value of .00. The (ns) stands for "not statistically significant" – in other words, not different from .00.

The important thing to remember is that levels of statistical significance (.05, .01, .001) only tell us how confident we are that the resulting sample correlation is truly different from a population correlation of zero. It does not mean that the correlation is high or that one value caused the other. With correlations, we can never assume that one measure causes another. Age and income are correlated but one would never dream of suggesting that with age you automatically become "well-to-do." Statistical significance is also affected by sample size. Even a low correlation of .006 can be highly significant if it is based upon an extremely large sample size. That outcome would merely tell the researcher that the population correlation is not likely to be .00. It is likely to be .006! The researcher would still be saddled with a very low correlation – incapable of predicting with it. Applied to assessment research, we should not think that there is a very useful consequence resulting from a very low correlation between test scores and grades even if the correlation is statistically significant. What counts in terms of usefulness is the magnitude of the correlation – not its significance level.

Collapsing Grade Cells: In most of the analyses that follow, the grade notations of D, F, NC, or WT (withdrew after census) are collapsed into one unsuccessful category and weighted as a "1" for numerical purposes. What we end up with are the following grades and weights: A = 4; B = 3; C or CR = 2; and D, F, NC, WT = 1. Weighing all grades in this way has proven very useful in terms of maintaining sample sizes, and has been referred to as the Research GPA. This scale also correlates with conventional GPA ($r = .90$) and success rate ($r = .90$).[1]

[1] Rasor, R. A. & Barr, J. (1993). Refinement in Assessment Validation: Technicalities of Dealing With Low Correlations and Instructor Grading Variation. American River College Research Report.

English Placement Test (Compass) for English 1A, 58, 256

English Writing Cut Scores

English Writing Placement	Writing Test Score
English 256	1 to 48
English 58	49 to 80
English 1A	81 to 98
English 1A Honors	99 to 100

ARC Normative Data

Sample size = 12,940 (all test records)
Maximum possible score = 100

Mean = 60.34
Standard Deviation = 29.30
Median = 65.00
Obtained Range = 1 to 99
Middle 50% Range = 35 to 87
Skewness = -.48 negative skew

Percentiles: 10th = 16, 20th = 28;
30th = 42; 40th = 55; 50th = 65;
60th = 74; 70th = 82; 80th = 90;
90th = 96

The data above are based upon the computerized Compass test, English scale. The data were generated by all students who took the test at American River College (or its satellites) irrespective of whether or not such students eventually enrolled in an ARC English course. The essential statistics for establishing local norms include the sample size, the mean and median, and the various percentile ranks. The mean (arithmetic average) is about 5 points lower than the median (middle point in a ranked distribution). This difference indicates that a collection of low scores "pulled down" the value of the mean whereas the median would be unaffected. This conclusion is also borne out by the degree of negative skewness, that is, the resulting histogram (curve) has more of a trail of very low scores than high scores. Very low scores are not too surprising in view of the fact that some students with a marginal ability to read English take the English scale only to stop after attempting a few items. However, we tried to eliminate all low scores that were a result of not proceeding correctly with a computer.

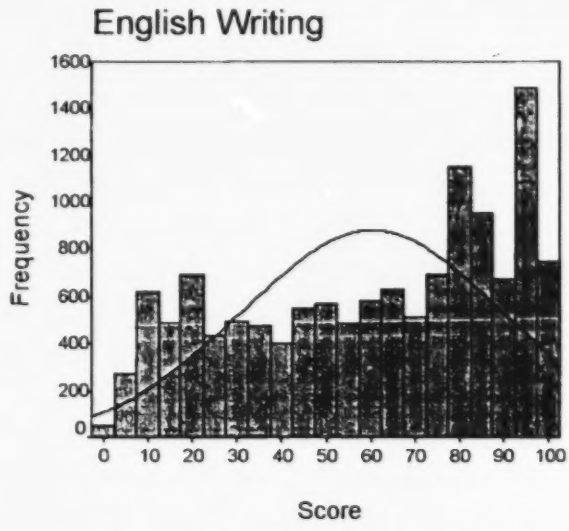

Figure 1. All Compass English Writing Assessment Scores

English 1A Analysis

Conditions for analysis: No evidence of enrolling in English 256 or 58 prior to 1A, test date prior to or beginning of English 1A, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 81 (68th percentile)	Total number of students this analysis: 1,175 (100%) Number of students below cut = 213 (18.1%) Number of students at/above cut = 962 (81.9%)

English 1A Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 63.40%	(b) 58.50%
Course Non-Success	(c) 36.60%	(d) 41.50%
Totals	100.00%	100.00%

English 1A Four-Cell Analysis
Overall Course Success Rate = 59.4%
Four-Cell Test Predictive Accuracy = 54.6% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.038 (ns)

English 1A Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	15.00%	23.40%
B	32.40%	24.80%
C or CR	16.00%	10.30%
D,F,NC,WT	36.60%	41.50%
Totals	100.00%	100.00%

Correlation Between Degrees of Success (1,2,3,4) and Cut Score Break (1,2) = .114 ($p<.01$)

English 1A Partial Range Grades and All Scores	
A = 4	Score Range
B = 3	English 1A = 5 to 99
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades and All Scores = .027 (ns)

Summary Comment on English 1A: When controlling for students' English course-taking history, the Compass English Test would seem to be a very weak instrument in terms of predicting final grades in English 1A. Even applying statistical manipulations to correct for such things as restricted range, grading variability, and deleting all data on students who withdrew, do not improve matters (see Table 1). All data configurations and correlational values are below what is acceptable for continued use of this test to place students into English 1A.

BEST COPY AVAILABLE

English 58 Analysis

Conditions for analysis: No evidence of enrolling in English 256 or 1A prior to 58, test date prior to or beginning of English 58, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 49 (36th percentile)	Total Number of students this analysis: 823 (100%) Number of students below cut = 188 (22.8%) Number of students at/above cut = 635 (77.2%)

English 58 Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 43.60%	(b) 47.70%
Course Non-Success	(c) 56.40%	(d) 52.30%
Totals	100.00%	100.00%

English 58 Four-Cell Analysis
Overall Course Success Rate = 46.8%
Four-Cell Test Predictive Accuracy Rate = 49.7% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .034 (ns)

English 58 Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	7.40%	15.90%
B	21.80%	21.90%
C or CR	14.40%	9.90%
D,F,NC,WT	56.40%	52.30%
Totals	100.00%	100.00%

Correlation Between Degrees of Success (1,2,3,4) and Cut Score Break (1,2) = .113 ($p<.05$)

English 58 Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	English 58 = 3 to 99
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades and All Scores = .111 ($p=.001$)

Summary Comment on English 58: When controlling for students' English course-taking history, the Compass English Test would seem to be a weak instrument in terms of predicting final grades in English 58. Although some of the correlation values are statistically significant from zero, they are still low. Applying statistical manipulations to correct for such things as restricted range, grading variability, and deleting all data on students who withdrew, did not improve the results to any satisfactory level (see Table 1). All data configurations and correlational values are below what is acceptable for continued use of this test to place students into English 58.

English 256 Analysis

Condition for analysis: No evidence of enrolling in English 58 or 1A prior to 256, test date prior to or beginning of English 256, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Former Cut Score = 16 (11th percentile)	Total Number of students this analysis: 419 (100%) Number of students below cut = 77 (18.4%) Number of students at/above cut = 342 (81.6%)

English 256 Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 46.80%	(b) 57.90%
Course Non-Success	(c) 53.20%	(d) 42.10%
Totals	100.00%	100.00%

English 256 Four-Cell Analysis
Overall Course Success Rate = 55.8%
Four-Cell Test Predictive Accuracy Rate = 57.0% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .087 (ns)

English 256 Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	6.50%	11.70%
B	22.10%	26.60%
C or CR	18.20%	19.60%
D,F,NC,WT	53.20%	42.10%
Totals	100.00%	100.00%

Correlation Between Degrees of Success (1,2,3,4) and Cut Score Break (1,2) = .097 (ns)

English 256 Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	English 256 = 1 to 84
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades and All Scores = .099 ($p<.05$)

Summary Comment on English 256: When applying controls for students' English course-taking history, the Compass English Test reveals some percentage differences in grades as related to cut score, but it still is a weak instrument in terms of predicting final grades in English 256. The application of statistical manipulations to correct for such things as restricted range, grading variability, and deleting all data on students who withdrew, did not improve the basic results (see Table 1). All data configurations and correlational values are below what is acceptable for continued use of this test to place students into English 256.

Reading Placement Test (Compass) for English 4, 5, 71 (now 78), & 268/270

English Reading Cut Scores

English Reading Placement	Reading Test Score
English 268 or 270	1 to 70
English 71 (78)	71 to 84
English 4 or 5	85 to 100

ARC Normative Data

Sample size = 13,304 (all test records)

Mean = 78.81
Standard Deviation = 16.47
Median = 83
Absolute Range = 19 to 99
Middle 50% Range = 70 to 92
Skewness = -.76 high negative skew

Percentiles: 10^{th} = 56; 20^{th} = 65;
30^{th} = 73; 40^{th} = 78; 50^{th} = 83;
60^{th} = 87; 70^{th} = 90; 80^{th} = 93; 90^{th} = 96

The data above are based upon the computerized Compass test, the scale for reading. The data were generated by all students who took the test at American River College (or its satellites) irrespective of whether or not such students eventually enrolled in an ARC reading course. The essential statistics for establishing local norms include the sample size, the mean and median, and the various percentile ranks. Like the English writing scores, these scores show a negative skewness, that is, the resulting histogram (curve) has more of a trail of very low scores than high scores. In discussions with the Compass testing research division, we found that they also found a negatively skewed curve when compiling data for community colleges across the nation.

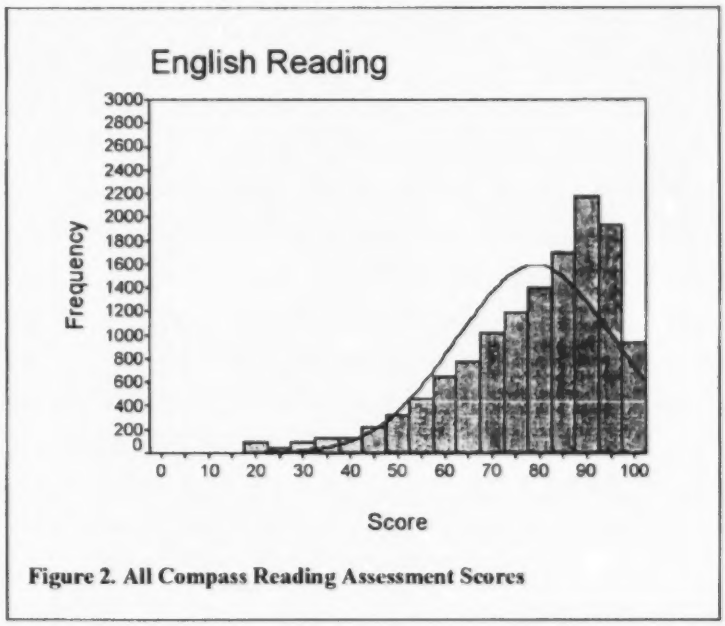

Figure 2. All Compass Reading Assessment Scores

English 4/5 Analysis

Conditions for analysis: No evidence of enrolling in English 268/270 or 71 prior to 4/5, test date prior to or beginning of English 4/5, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 85 (57th percentile)	Total number of students this analysis: 72 (100%) Number of students below cut = 16 (22.2%) Number of students at/above cut = 56 (77.8%)

English 4/5 Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 62.50%	(b) 37.50%
Course Non-Success	(c) 37.50%	(d) 62.50%
Totals	100.00%	100.00%

English 4/5 Four-Cell Analysis
Overall Course Success Rate = 43.1%
Four-Cell Test Prediction Accuracy = 37.5% (b + c cell frequencies relative to total sample size
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.210 (ns)
Caution: Sample size relatively small for definitive interpretation. Consider only as trend.

English 4/5 Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	12.50%	10.70%
B	31.30%	14.30%
C or CR	18.80%	12.50%
D,F,NC,WT	37.50%	62.50%
Totals	100.00%	100.00%

Correlation Between Degrees of Success (1,2,3,4) and Cut Score Break (1,2) = -.228 (ns)

Caution: Sample size relatively small for definitive interpretation.

English 4/5 Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	English 4/5 = 36 to 99
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades and All Scores = -.015 (ns)
Caution: Sample size relatively small for definitive interpretation. Consider only as trend.

Summary Comment on English 4/5: With only 72 students in this analysis, little can be said other than the negative signed correlations are alarming because students who fell below the cut score experienced higher success than students scoring above the cut score. Should this trend continue with larger sample sizes, immediate discontinuance of the test would be recommended.

BEST COPY AVAILABLE

English 71 Analysis (Course number has now been changed to 78)

Conditions for analysis: No evidence of enrolling in English 268/270 prior to 71, test date prior to or beginning of English 71, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 71 (28th percentile)	Total number of students this analysis: 266 (100%) Number of students below cut = 26 (9.8%) Number of students at/above cut = 240 (90.2%)

English 71 Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 61.50%	(b) 47.90%
Course Non-Success	(c) 38.50%	(d) 52.10%
Totals	100.00%	100.00%

English 71 Four-Cell Analysis
Overall Course Success Rate = 49.2%
Four-Cell Test Prediction Accuracy = 47.0% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.081 (ns)

English 71 Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	34.60%	20.00%
B	23.10%	15.40%
C or CR	3.80%	12.50%
D,F,NC,WT	38.50%	52.10%
Totals	100.00%	100.00%

Correlation Between Degrees of Success (1,2,3,4) and Cut Score Break (1,2) = -.145 (ns)

English 71 Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	English 71 = 58 to 99
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades and All Scores = -.133 ($p<.05$ but in wrong direction)

Summary Comment on English 71 (now 78): There were only 26 students in this analysis who had a score below the cut point. This puts a limit on what can be said about the suitability of the cut score being 71. However, the correlation between grades and all scores (cut score not a factor) is not only low but in the wrong direction (a negative). This means that there is a tendency for lower scoring students to earn higher grades, while higher scoring students earn lower grades. This last result, which is based upon 266 students, would suggest immediate discontinuance of the test as used with English 78.

English 268/270 Analysis

Conditions for analysis: No evidence of enrolling in English 71 prior to 268/270, test date prior to or beginning of English 268/270, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 53 (8th percentile)	Total number of students this analysis: 461 (100%) Number of students below cut = 108 (23.4%) Number of students at/above cut = 353 (76.6%)

English 268/270 Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 36.10%	(b) 49.30%
Course Non-Success	(c) 63.90%	(d) 50.70%
Totals	100.00%	100.00%

English 268/270 Four-Cell Analysis
Overall Course Success Rate = 46.2%
Four-Cell Test Prediction Accuracy = 52.7% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .112 ($p<.05$)

English 268/270 Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	5.60%	16.40%
B	5.60%	11.90%
C or CR	25.00%	21.00%
D,F,NC,WT	63.90%	50.70%
Totals	100.00%	100.00%

Correlation Between Degrees of Success (1,2,3,4) and Cut Score Break (1,2) = .171 ($p<.01$)

English 268/270 Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	English 268/270 = 19 to 99
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades and All Scores = .131 ($p<.01$)

Summary Comment on English 268/270: All of the correlations between scores and grades are in the positive direction, significantly different from zero, but still quite low. For this reason, the Compass test is not recommended for placement in English 268/270.

Math Placement Test (Compass) for Math 51, 53, 15

There are five mathematical test levels within the Compass instrument. Cut scores established for entry into a particular math course are based upon one of two scales, e.g., a relative high score on a lower level math test, or a relatively low score on a next higher test.

Math Cut Scores

Math Placement	Math Test & Score
Math 51 (Algebra 1)	Math 1: 48 to 100
Math 51 (Algebra 1)	Math 2: 26 to 40
Math 53 (Algebra 2)	Math 2: 41 to 64
Math 53 (Algebra 2)	Math 3: 31 to 49
Math 15 (Trig.)	Math 4: 52 to 100
Math 15 (Trig.)	Math 5: 31 to 40

ARC Normative Data for Math Placement (Compass)

Math 1 Compass Test

Sample size = 8,206 (all test records)
Mean = 35.15
Standard Deviation = 14.91

Median = 33
Absolute Range = 17 to 99
Middle 50% Range = 24 to 41
Skewness = .433

Percentiles: 10^{th} = 20; 20^{th} =23;
30^{th} = 26; 40^{th} = 29; 50^{th} = 33;
60^{th} = 36; 70^{th} = 39; 80^{th} = 44;
90^{th} = 52

Math 2 Compass Test

Sample size = 3,136 (all test records)
Mean = 40.39
Standard Deviation = 12.83

Median = 38
Absolute Range = 15 to 96
Middle 50% Range = 30 to 50
Skewness = .559

Percentiles: 10^{th} = 26; 20^{th} = 29;
30^{th} = 31; 40^{th} = 34; 50^{th} = 38;
60^{th} = 42; 70^{th} = 47; 80^{th} = 53;
90^{th} = 59

Math 3 Compass Test

Sample size = 310 (all test records)
Mean = 42.08
Standard Deviation = 7.62

Median = 42
Absolute Range = 25 to 92
Middle 50% Range = 36 to 47
Skewness = .031

Percentiles: 10^{th} = 33; 20^{th} = 35;
30^{th} = 38; 40^{th} = 40; 50^{th} = 42;
60^{th} = 44; 70^{th} = 46; 80^{th} = 47;
90^{th} = 49

Math 4 Compass Test

Sample size = 282 (all test records)
Mean = 67.16
Standard Deviation = 17.23

Median = 68
Absolute Range = 31 to 99
Middle 50% Range = 53 to 81
Skewness = -.146

Percentiles: 10^{th} = 43; 20^{th} = 50;
30^{th} = 56; 40^{th} = 64; 50^{th} = 68;
60^{th} = 72; 70^{th} = 78; 80^{th} = 84;
90^{th} = 91

ARC Normative Data for Math Placement (Compass) Continued

Math 5 Compass Test

Sample size = 817 (all test records)
Mean = 54.59
Standard Deviation = 15.13

Median = 52
Absolute Range = 31 to 99
Middle 50% Range = 44 to 64
Skewness = .514

Percentiles: 10^{th} = 37; 20^{th} = 41;
30^{th} = 45; 40^{th} = 49; 50^{th} = 52;
60^{th} = 56; 70^{th} = 60; 80^{th} = 67; 80^{th} = 67;
90^{th} = 75

The sample sizes for math tests 3, 4, and 5 are somewhat small for establishing definitive ARC norms. During the next period for publishing "Institutional Effectiveness," the sample sizes will have increased sufficiently thereby adding more confidence in any statistical analyses using these tests. In the outcomes described next, we combined the two groups who took different level math tests but enrolled in the same course. Individuals in one of the two groups either: 1) scored below the designated cut score on either test appropriate for the specific course, or 2) scored at/above the cut score on either test. For example, cut scores for the math 51 course: Math1 test = 48 or math2 test =26.

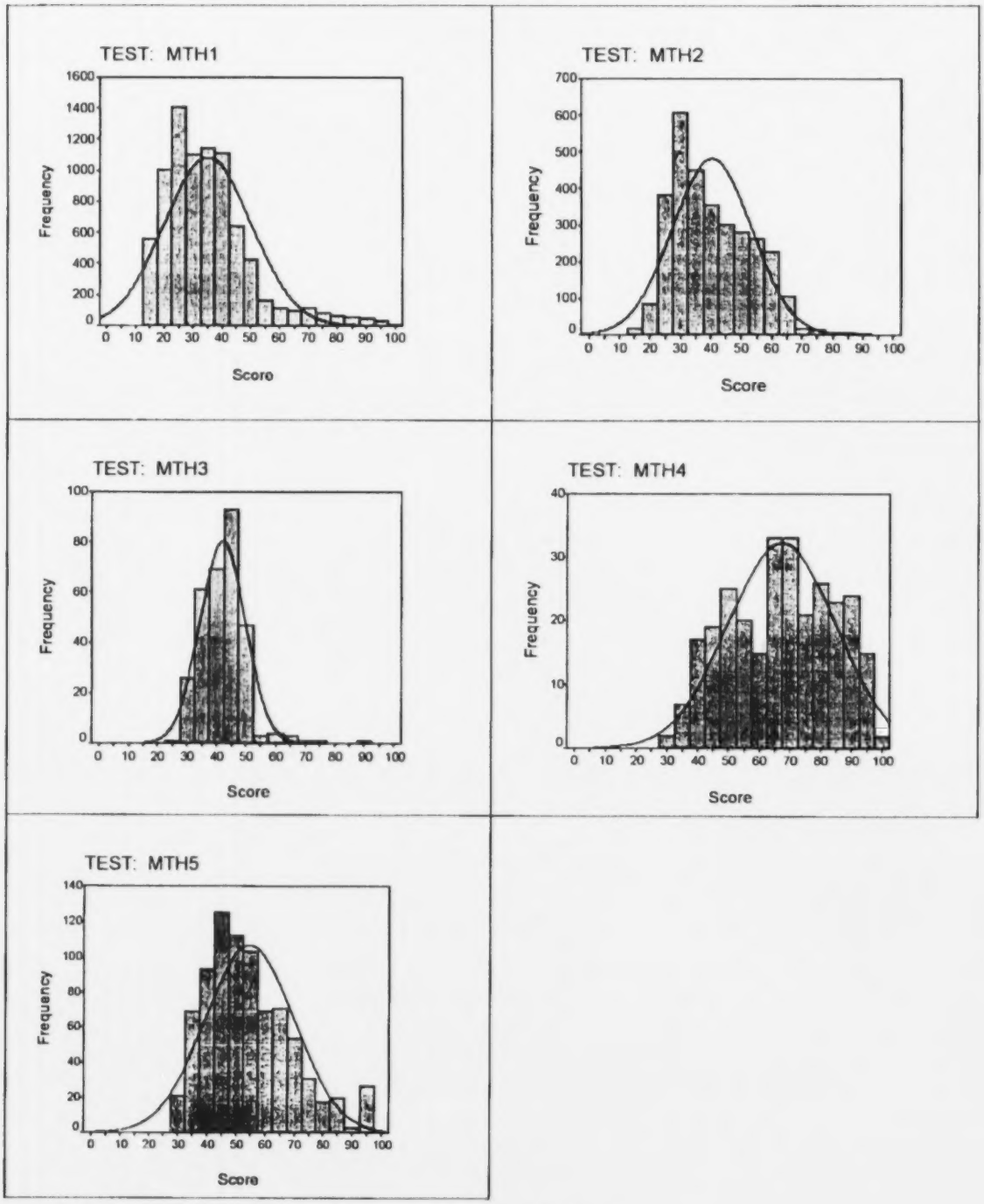

Figure 3. ARC Normative Distributions for Five Compass Math Tests.

Math 51 Analysis

Conditions for analysis: No evidence of enrolling in a lower level math course, test date prior to or beginning of Math 51, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 48 on Math 1 Test (87th percentile) or 26 on Math 2 Test (11th percentile)	Total number of students this analysis: 723 (100%) Number of students below cut = 289 (40.0%) Number of students at/above cut = 434 (60.0%)

Math 51 Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 41.20%	(b) 58.10%
Course Non-Success	(c) 58.80%	(d) 41.90%
Totals	100.00%	100.00%

Math 51 Four-Cell Analysis
Overall Course Success Rate = 51.3%
Four-Cell Test Prediction Accuracy = 58.4% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) & Cut Score Break (1,2) = .166 (p<.001)

Math 51 Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	14.20%	22.60%
B	12.50%	20.00%
C or CR	14.50%	15.40%
D,F,NC,WT	58.80%	41.90%
Totals	100.00%	100.00%

Correlation Between Degrees of Success (1,2,3,4) & Cut Score Break (1,2) = .176 (p<.001)

Math 51 Partial Range Grades and Scores	
A = 4	Score Range
B = 3	Math 1 = 32 to 95
C or CR = 2	Math 2 = 15 to 64
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) & all Math 1 Test Scores (n = 441) = .173 (p<.001)

Correlation Between Grades (1,2,3,4) & all Math 2 Test Scores (n = 282) = .235 (p<.001)

Summary Comment on Math 51: All correlation values between test cut-off levels and grades are statistically different from zero, yet they are low – too low to suggest anything like accurate prediction of grade or success from knowledge of test score. However, the percentage differences (above vs. below cut) and various grades show that the test is predicting in the correct direction, i.e., higher percentage of A's, B's, and C's in the at/above cut score group. When statistical manipulations were made to adjust for such things as restricted range, grading variability among instructors, and deleting data on students who withdrew from the course, correlation values did not change appreciably (see Table 1). In view of the present results, the use of the Compass math scales for entrance into Math 51 cannot be given a strong endorsement. Generally, the lowest acceptable level of correlation for situations like this is .36 which accounts for 13% of the variance in grades.

BEST COPY AVAILABLE

Math 53 Analysis

Conditions for analysis: No evidence of enrolling in Math 51 or a lower level math course, test date prior to or beginning of Math 53, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 41 on Math 2 Test (59th percentile) or 31 on Math 3 Test (5th percentile)	Total number of students this analysis: 324 (100%) Number of students below cut = 92 (28.4%) Number of students at/above cut = 232 (71.6%)

Math 53 Four-Cell Analysis		
	Below	At/Above
Outcome	Cut Score	Cut Score
Course Success	(a) 47.80%	(b) 62.90%
Course Non-Success	(c) 52.20%	(d) 37.10%
Totals	100.00%	100.00%

Math 53 Four-Cell Analysis
Overall Course Success Rate = 58.6%
Four-Cell Test Prediction Accuracy = 59.9% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) & Cut Score Break (1,2) = .138 (p <.05)

Math 53 Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	10.90%	20.70%
B	20.70%	20.70%
C or CR	16.30%	21.60%
D,F,NC,WT	52.20%	37.10%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) & Cut Score Break (1,2) = .158 (p<.05)

Math 53 Partial Range Grades and Scores	
A = 4	Score Range
B = 3	Math 2 = 21 to 82
C or CR = 2	Math 3 = 28 to 49
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) & all Math 2 Test Scores (n=289) = .198 (p=.001)
Correlation Between Grades (1,2,3,4) & all Math 3 Test Scores (n=35) = .054 (ns)

Summary Comment on Math 53: As with Math 51, nearly all correlation values between test cut-off levels and grades are statistically different from zero, yet they are also low – too low to enable accurate prediction of grade or success from knowledge of test score. It is acknowledged that the sample sizes are small. Only time will tell if the correlation values increase with larger sample sizes. The percentage differences (above vs. below cut) and various grades show that the test is at least predicting in the correct direction, i.e., higher percentage of A's, B's, and C's in the at/above cut score group. When statistical adjustments were made for such things as restricted range, grading variability among instructors, and deleting data on students who withdrew from the course, correlation values did not change for the better (see Table 1). Given these results, the use of the Compass math scales for entrance into Math 53 is not encouraged.

Math 15 Analysis

Conditions for analysis: No evidence of enrolling in Math 53 or a lower level math course, test date prior to or beginning of Math 15, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 52 on Math 4 Test (25th percentile) or 31 on Math 5 Test (1st percentile)	Total number of students this analysis: 50 (100%) Number of students below cut = 3 (6.0%) Number of students at/above cut = 47 (94.0%)

Math 15 Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 66.70%	(b) 70.20%
Course Non-Success	(c) 33.30%	(d) 29.80%
Totals	100.00%	100.00%

Math 15 Four-Cell Analysis
Overall Course Success Rate = 70.0%
Four-Cell Test Prediction Accuracy = 68.0% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) & Cut Score Break (1,2) = -.018 (ns)
Caution: Sample size too small for definitive interpretation. Consider only as trend.

Math 15 Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	33.30%	34.00%
B	33.30%	21.30%
C or CR	0.00%	14.90%
D,F,NC,WT	33.30%	29.80%
Totals	100.00%	100.00%

Correlation Between Degrees of Success (1,2,3,4) & Cut Score Break (1,2) = -.114 (ns)

Caution: Sample size too small for definitive interpretation. Consider only as trend.

Math 15 Partial Range Grades and Scores	
A = 4	Score Range
B = 3	Math 4 = 48 to 94
C or CR = 2	Math 5 = 32 to 59
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) & all Math 4 Test Scores (n=28) = -.022 (ns)

Correlation Between Grades (1,2,3,4) & all Math 5 Test Scores (n=22) = .065 (ns)

Caution: Sample size relatively small for definitive interpretation. Consider only as trend.

Summary Comment on Math 15: Unfortunately, no definitive conclusions can be drawn due to low sample sizes. However, what data exist about the relationship between Compass tests and the trigonometry course is not very encouraging in so far as using the test for course placement.

ESL Placement Test (Compass) for ESL 260, 270, 280, 1, 2, 63, 5

ESL Cut Scores

ESL Placement	Reading	Writing	Listening
See Counselor	1 to 27	1 to 27	1 to 27
ESL 260	28 to 37	28 to 41	28 to 41
ESL 270	38 to 51	42 to 55	42 to 59
ESL 280	52 to 64	56 to 65	60 to 66
ESL 1	65 to 72	66 to 75	67 to 74
ESL 2	73 to 80	76 to 86	75 to 100
ESL 63	81 to 94	87 to 94	
ESL 5	95 to 100	95 to 100	

ARC Normative Data

ESL Reading (ESLR)

Sample size = 4,028 (all test records)
Mean = 69.70
Standard Deviation = 17.53

Median = 74.00
Absolute Range = 25 to 99
Middle 50% Range = 57 to 84
Skewness = -.736 high negative skew

Percentiles: 10^{th} = 43; 20^{th} = 53; 30^{th} = 61; 40^{th} = 68; 50^{th} =74; 60^{th} = 78; 70^{th} = 82; 80^{th} =86; 90^{th} = 90

ESL Writing (ESLW)

Sample size = 3,998 (all test records)
Mean = 63.56
Standard Deviation = 15.50

Median = 64
Absolute Range = 25 to 99
Middle 50% Range = 53 to 76
Skewness = -.085

Percentiles: 10^{th} = 42; 20^{th} = 49; 30^{th} = 55; 40^{th} = 60; 50^{th} = 64; 60^{th} = 69; 70^{th} = 74; 80^{th} = 78; 90^{th} = 83

ESL Listening (ESLL)

Sample size = 3,919
Mean = 66.42
Standard Deviation = 17.31

Median = 69
Absolute Range = 24 to 99
Middle 50% Range = 54 to 79
Skewness = -.447

Percentiles: 10^{th} = 42; 20^{th} = 52; 30^{th} = 58; 40^{th} = 63; 50^{th} = 69; 60^{th} = 73; 70^{th} = 77; 80^{th} = 82; 90^{th} = 88

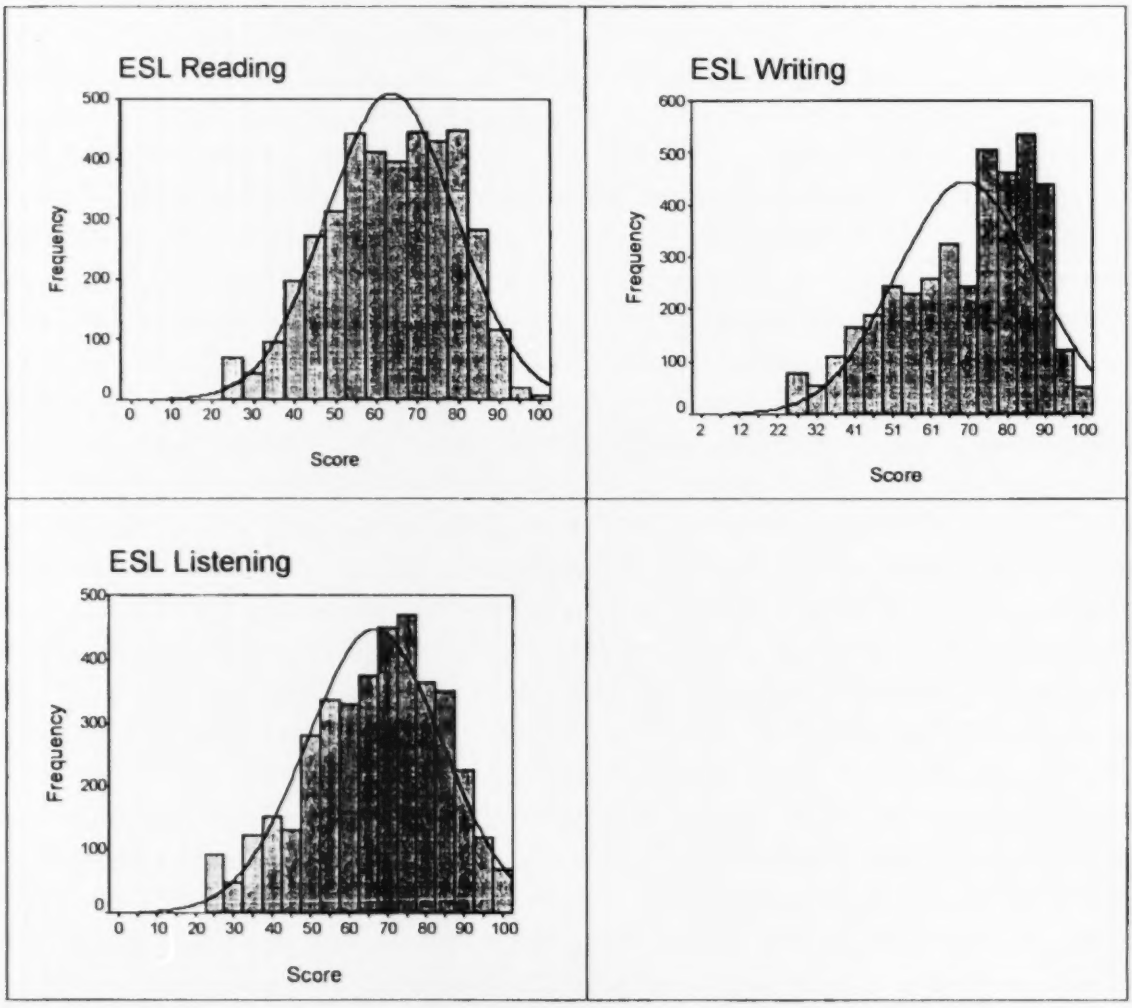

Figure 4. ARC Normative Distributions for Three Compass ESL Tests.

Eight Levels of ESL Course Placement with Percentiles

Course	ESLR Score	ESLR Percentile		ESLW Score	ESLW Percentile		ESLL Score	ESLL Percentile
1 = ESL 5	95	98th		95	99.6th		75	67th
2 = ESL 63	81	70th		87	96th		75	67th
3 = ESL 2	73	49th		76	76th		75	67th
4 = ESL 1	65	36th		66	54th		67	48th
5 = ESL 280	52	19th		56	32nd		60	36th
6 = ESL 270	38	6th		42	10th		42	11th
7 = ESL 260	28	2nd		28	2nd		28	3rd
8 = See a Counselor	1	1st		1	1st		1	1st

Recently Implemented: Seven Levels of ESL Course Placement with Percentiles

Course	ESLR Score	ESLR Percentile		ESLW Score	ESLW Percentile		ESLL Score	ESLL Percentile
1 = ESL 5/63	95	98th		95	99.6th		91	95th
3 = ESL 2	87	85th		87	96th		91	95th
4 = ESL 1	76	56th		76	76th		75	67th
5 = ESL 280	66	37th		66	54th		67	48th
6 = ESL 270	56	25th		56	32nd		60	36th
7 = ESL 260	31	3rd		31	3rd		31	3rd
8 = See a Counselor	1	1st		1	1st		1	1st

BEST COPY AVAILABLE

ESL 2 R (Reading) Analysis

Conditions for analysis: No evidence of enrolling in any lower level ESL reading course, test date prior to or beginning of ESL 2R, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 73 on ESLR Test (49th percentile)	Total number of students this analysis: 44 (100%) Number of students below cut = 2 (4.5%) Number of students at/above cut = 42 (95.5%)

ESL 2R Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 100.00%	(b) 88.10%
Course Non-Success	(c) 0.00%	(d) 11.90%
Totals	100.00%	100.00%

ESL 2Rfour-Cell Analysis
Overall Course Success Rate = 88.6%
Four-Cell Test Prediction Accuracy = 84.0% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.078 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 2R Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	50.00%	59.50%
B	50.00%	14.30%
C or CR	0.00%	14.30%
D,F,NC,WT	0.00%	11.90%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = -.218 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 2R Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 2R = 48 to 99
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 2R Test Scores = .370 ($p<.05$)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 2R: With only a sample size of 44 students, very little can be said of the analysis other than the overall correlation between all grades and all scores appears promising in terms of being useful for course placement. See Table 2 for summary.

ESL 1 R (Reading) Analysis

Condition for analysis: No evidence of enrolling in any lower level ESL reading course, test date prior to or beginning of ESL 1R, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 65 on ESLR Test (36th percentile)	Total number of students this analysis: 21(100%) Number of students below cut = 2 (9.5%) Number of students at/above cut = 19 (90.5%)

ESL 1R Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 100.00%	(b) 57.90%
Course Non-Success	(c) 0.00%	(d) 42.10%
Totals	100.00%	100.00%

ESL 1R Four-Cell Analysis
Overall Course Success Rate = 61.9%
Four-Cell Test Prediction Accuracy = 52.4% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.255 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 1R Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	0.00%	10.50%
B	50.00%	36.80%
C or CR	50.00%	10.50%
D,F,NC,WT	0.00%	42.10%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = -.385 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 1R Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 1R = 48 to 78
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 1R Test Scores = -.207 (ns)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 1R: With a small sample size of 21 students, little can be said of the various analyses. The preliminary results with negative correlations are not at all encouraging for using this test with placement into ESL 1R. See Table 2 for summary.

BEST COPY AVAILABLE

ESL 280 R (Reading) Analysis

Condition for analysis: No evidence of enrolling in any lower level ESL reading course, test date prior to or beginning of ESL 280 R, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 52 on ESLR Test (19th percentile)	Total number of students this analysis: 50 (100%) Number of students below cut = 2 (4.0%) Number of students at/above cut = 48 (96.0%)

ESL 280 R Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 50.00%	(b) 77.10%
Course Non-Success	(c) 50.00%	(d) 22.90%
Totals	100.00%	100.00%

ESL 280 R Four-Cell Analysis
Overall Course Success Rate = 76.0%
Four-Cell Test Prediction Accuracy = 76.0% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .124 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 280 R Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	0.00%	22.90%
B	50.00%	12.50%
C or CR	0.00%	41.70%
D,F,NC,WT	50.00%	22.90%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .276 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 280 R Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 280R = 41 to 95
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 280R Test Scores = .227 (ns)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 280R: The sample size of 50 students and with only 2 falling below the cut score precludes any definitive conclusion. The preliminary results between all scores and grades (.227) is still too small in value to effectively place students into ESL 280R. See Table 2 for summary.

ESL 270 R (Reading) Analysis

Condition for analysis: No evidence of enrolling in any lower level ESL reading course, test date prior to or beginning of ESL 270 R, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 38 on ESLR Test (6th percentile)	Total number of students this analysis: 68 (100%) Number of students below cut = 0 (0.0%) Number of students at/above cut = 68 (100.0%)

ESL 270 R Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 00.00%	(b) 85.30%
Course Non-Success	(c) 00.00%	(d) 14.70%
Totals	100.00%	100.00%

ESL 270 R Four-Cell Analysis
Overall Course Success Rate = 85.3%
Four-Cell Test Prediction Accuracy = 85.3% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .00 (ns)
Caution: Sample size too small and no cases below the cut prevents interpretation.

ESL 270 R Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	0.00%	47.10%
B	0.00%	22.10%
C or CR	0.00%	16.20%
D,F,NC,WT	0.00%	14.70%
Totals	00.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .00 (ns)

Caution: Sample size too small and no cases below the cut prevents interpretation.

ESL 270 R Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 270R = 38 to 91
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 270R Test Scores =.425 (p< .001)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 270R: The sample size of 68 students and no one falling below the cut score prevents any definitive conclusion. However, the preliminary correlation between all grades and all scores is encouraging. Such a finding on a larger sample would suggest that a meaningful cut score could be found. See Table 2 for summary.

BEST COPY AVAILABLE

ESL 260 R (Reading) Analysis

Condition for analysis: Test date prior to or beginning of ESL 260 R, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 28 on ESLR Test (2nd percentile)	Total number of students this analysis: 133 (100%) Number of students below cut = 14 (10.5%) Number of students at/above cut = 119 (89.5%)

ESL 260 R Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 21.40%	(b) 73.90%
Course Non-Success	(c) 78.60%	(d) 26.10%
Totals	100.00%	100.00%

ESL 260 R Four-Cell Analysis
Overall Course Success Rate = 68.4%
Four-Cell Test Prediction Accuracy = 74.4% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .347 (p<.001)

ESL 260 R Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	0.00%	0.00%
B	0.00%	0.00%
C or CR	21.40%	73.90%
D,F,NC,WT	78.60%	26.10%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = (see 4-cell outcome)

Caution: Course graded on Credit/No Credit basis. See Four-Cell analysis.

ESL 260 R Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 260R = 25 to 80
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Credit/No Credit Grades (1,2) and All 260R Test Scores = .255 (p<.01)

Summary Comment on ESL 260R: The sample size of 119 students is barely acceptable for analysis. The preliminary correlations (four-cell and full range between credit/no credit and all scores) suggest that given a larger sample size, the Compass test may approach usefulness for placement in this course. See Table 2 for summary.

ESL 2 W (Writing) Analysis

Condition for analysis: No evidence of enrolling in any lower level ESL writing course, test date prior to or beginning of ESL 2W, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 76 on ESLW Test (76th percentile)	Total number of students this analysis: 41 (100%) Number of students below cut = 21 (51.2%) Number of students at/above cut = 20 (48.8%)

ESL 2W Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 38.10%	(b) 40.00%
Course Non-Success	(c) 61.90%	(d) 60.00%
Totals	100.00%	100.00%

ESL 2W Four-Cell Analysis
Overall Course Success Rate = 39.0%
Overall Course Success Rate = 39.0%
Four-Cell Test Prediction Accuracy = 51.2% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .020 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 2W Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	4.80%	10.00%
B	19.00%	20.00%
C or CR	14.30%	10.00%
D,F,NC,WT	61.90%	60.00%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .116 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 2W Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 2W = 57 to 94
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 2W Test Scores = .098 (ns)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 2W: With a sample size of only 41 students, very little can be said of the various analyses other than the preliminary results are not encouraging in so far as using this test for placement into ESL 2W. See Table 2 for summary.

ESL 1 W (Writing) Analysis

Condition for analysis: No evidence of enrolling in any lower level ESL writing course, test date prior to or beginning of ESL 1W, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 66 on ESLW Test (54th percentile)	Total number of students this analysis: 28 (100%) Number of students below cut = 6 (21.4%) Number of students at/above cut = 22 (78.6%)

ESL 1W Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 66.70%	(b) 45.00%
Course Non-Success	(c) 33.30%	(d) 54.50%
Totals	100.00%	100.00%

ESL 1W Four-Cell Analysis
Overall Course Success Rate = 50.0%
Four-Cell Test Prediction Accuracy = 42.9% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.174 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 1W Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	0.00%	13.60%
B	16.70%	22.70%
C or CR	50.00%	9.10%
D,F,NC,WT	33.30%	54.50%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .453 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 1W Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 1W = 47 to 84
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 1W Test Scores = -.028 (ns)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 1W: While the sample size is too small for reliable conclusions, these early results are not very encouraging in so far as using this test for placement into ESL 1W. See Table 2 for summary.

BEST COPY AVAILABLE

ESL 280 W (Writing) Analysis

Condition for analysis: No evidence of enrolling in any lower level ESL writing course, test date prior to or beginning of ESL 280W, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 56 on ESLW Test (32nd percentile)	Total number of students this analysis: 48 (100%) Number of students below cut = 14 (29.2%) Number of students at/above cut = 34 (70.8%)

ESL 280 W Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 71.40%	(b) 70.60%
Course Non-Success	(c) 28.60%	(d) 29.40%
Totals	100.00%	100.00%

ESL 280 W Four-Cell Analysis
Overall Course Success Rate = 70.8%
Four-Cell Test Prediction Accuracy = 58.3% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.008 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 280 W Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	7.10%	5.90%
B	14.30%	29.40%
C or CR	50.00%	35.30%
D,F,NC,WT	28.60%	29.40%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .175 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 280 W Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 280 W = 41 to 86
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 280 W Test Scores = .118 (ns)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 280 W: The sample size (n = 48) is too small for reliable conclusions. Preliminary results do not point to accurate prediction. See Table 2 for summary.

ESL 270 W (Writing) Analysis

Conditions for analysis: No evidence of enrolling in any lower level ESL writing course, test date prior to or beginning of ESL 270W, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 42 on ESLW Test (10th percentile)	Total number of students this analysis: 64 (100%) Number of students below cut = 5 (7.8%) Number of students at/above cut = 59 (92.2%)

ESL 270 W Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 80.00%	(b) 74.60%
Course Non-Success	(c) 20.00%	(d) 25.40%
Totals	100.00%	100.00%

ESL 270 W Four-Cell Analysis
Overall Course Success Rate = 75.0%
Four-Cell Test Prediction Accuracy = 70.3% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.034 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 270 W Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	0.00%	10.20%
B	20.00%	30.50%
C or CR	60.00%	33.90%
D,F,NC,WT	20.00%	25.40%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .159 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 270 W Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 270 W = 29 to 82
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 270 W Test Scores = .298 ($p < .05$)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 270 W: The sample size (n = 64) is too small for a reliable conclusion. Existing correlations are low. See Table 2 for summary.

ESL 260 W (Writing) Analysis

Conditions for analysis: Test date prior to or beginning of ESL 260W, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 28 on ESLW Test (2nd percentile)	Total number of students this analysis: 150 (100%) Number of students below cut = 12 (8.0%) Number of students at/above cut = 138 (92.0%)

ESL 260 W Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 16.70%	(b) 62.30%
Course Non-Success	(c) 83.30%	(d) 37.70%
Totals	100.00%	100.00%

ESL 260 W Four-Cell Analysis
Overall Course Success Rate = 58.7%
Four-Cell Test Prediction Accuracy = 64.0% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .252 (p<.01)
Caution: Sample size in below cut group too small for definitive interpretation.

ESL 260 W Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	0.00%	0.00%
B	0.00%	0.00%
C or CR	16.70%	62.30%
D,F,NC,WT	83.30%	37.70%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = (see four-cell outcome)

Caution: Sample size is small for the below cut group (n=12). Course graded on credit/no credit basis.

ESL 260 W Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 260 W = 25 to 79
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Credit/No credit Grades (1,2) and All 260 W Test Scores = .402 (p<.001)

Summary Comment on ESL 260 W: The total sample size is 150 students. However, the below cut group only has 12 students. The existing correlation based upon a full range of scores and no cut score (.402), suggests that this test may be useful for placement into ESL 4W. See Table 2 for summary.

ESL 2 L (Listening) Analysis

Conditions for analysis: No evidence of enrolling in any lower level ESL listening course, test date prior to or beginning of ESL 2L, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 75 on ESLL Test (67th percentile)	Total number of students this analysis: 46 (100%) Number of students below cut = 6 (13.0%) Number of students at/above cut = 40 (87.0%)

ESL 2L Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 83.30%	(b) 67.50%
Course Non-Success	(c) 16.70%	(d) 32.50%
Totals	100.00%	100.00%

ESL 2L Four-Cell Analysis
Overall Course Success Rate = 69.6%
Four-Cell Test Prediction Accuracy = 60.9% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.116 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 2L Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	50.00%	30.00%
B	33.30%	32.50%
C or CR	0.00%	5.00%
D,F,NC,WT	16.70%	32.50%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = -.173 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 2L Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 2L = 34 to 99
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 2L Test Scores = .153 (ns)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 2L: With a sample size of only 46 students, little can be said of the various analyses. Preliminary results are not encouraging in so far as using this test for placement into ESL 2L. See Table 2 for summary.

ESL 1 L (Listening) Analysis

Conditions for analysis: No evidence of enrolling in any lower level ESL listening course, test date prior to or beginning of ESL 1L, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 67 on ESLL Test (48th percentile)	Total number of students this analysis: 50 (100%) Number of students below cut = 10 (20.0%) Number of students at/above cut = 40 (80.0%)

ESL 1L Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 80.00%	(b) 80.00%
Course Non-Success	(c) 20.00%	(d) 20.00%
Totals	100.00%	100.00%

ESL 1L Four-Cell Analysis
Overall Course Success Rate = 80.0%
Four-Cell Test Prediction Accuracy = 68.0% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .000 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 1L Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	30.00%	25.00%
B	20.00%	40.00%
C or CR	30.00%	15.00%
D,F,NC,WT	20.00%	20.00%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .198 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 1L Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 1L = 40 to 99
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 1L Test Scores = .051 (ns)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 1L: With a sample size of only 50 students, little can be said of the various analyses. Preliminary results are not encouraging for using this test with placement into ESL 1L. See Table 2 for summary.

BEST COPY AVAILABLE

ESL 280 L (Listening) Analysis

Conditions for analysis: No evidence of enrolling in any lower level ESL listening course, test date prior to or beginning of ESL 280 L, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 60 on ESLL Test (36th percentile)	Total number of students this analysis: 56 (100%) Number of students below cut = 15 (26.8%) Number of students at/above cut = 41 (73.2%)

ESL 280 L Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 80.00%	(b) 85.40%
Course Non-Success	(c) 20.00%	(d) 14.60%
Totals	100.00%	100.00%

ESL 280L Four-Cell Analysis
Overall Course Success Rate = 83.9%
Four-Cell Test Prediction Accuracy = 67.9% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .065 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 280 L Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	20.00%	26.80%
B	26.70%	22.00%
C or CR	33.30%	36.60%
D,F,NC,WT	20.00%	14.60%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .098 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 280 L Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 280L = 24 to 89
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 280L Test Scores = .277 ($p<.05$)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 280L: With a sample size of only 56 students, little can be said of the various analyses. Preliminary results are not encouraging in so far as use of this test with placing students into ESL 280L. See Table 2 for summary.

ESL 270 L (Listening) Analysis

Conditions for analysis: No evidence of enrolling in any lower level ESL listening course, test date prior to or beginning of ESL 270 L, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 42 on ESLL Test (11th percentile)	Total number of students this analysis: 65 (100%) Number of students below cut = 6 (9.2%) Number of students at/above cut = 59 (90.8%)

ESL 270 L Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 83.30%	(b) 76.30%
Course Non-Success	(c) 16.70%	(d) 23.70%
Totals	100.00%	100.00%

ESL 270L Four-Cell Analysis
Overall Course Success Rate = 76.9%
Four-Cell Test Prediction Accuracy = 70.8% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = -.049 (ns)
Caution: Sample size too small for definitive interpretation.

ESL 270 L Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	33.30%	28.80%
B	16.70%	37.30%
C or CR	33.30%	10.20%
D,F,NC,WT	16.70%	23.70%
Totals	100.00%	100.00%

Correlation Between Success (1,2,3,4) and Cut Score Break (1,2) = .221 (ns)

Caution: Sample size too small for definitive interpretation.

ESL 270 L Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 270L = 24 to 79
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Grades (1,2,3,4) and All 270L Test Scores = .269 ($p<.05$)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 270L: With a sample size of 65 students and only 6 falling below the cut score, few conclusions can be made. Preliminary results show low correlations. See Table 2 for summary.

ESL 260 L (Listening) Analysis

Conditions for analysis: Test date prior to or beginning of ESL 260 L, and highest score of record for student. Course success is defined as the percentage of final grade notations that are A, B, C, or CR.

Cut Score = 28on ESLL Test (3rd percentile)	Total number of students this analysis: 115 (100%) Number of students below cut = 13 (11.3%) Number of students at/above cut = 102 (88.7%)

ESL 260 L Four-Cell Analysis		
Outcome	Below Cut Score	At/Above Cut Score
Course Success	(a) 61.50%	(b) 74.50%
Course Non-Success	(c) 38.50%	(d) 25.50%
Totals	100.00%	100.00%

ESL 260L Four-Cell Analysis
Overall Course Success Rate = 73.0%
Four-Cell Test Prediction Accuracy = 70.4% (b + c cell frequencies relative to total sample size)
Correlation Between Success (1,2) and Cut Score Break (1,2) = .093 (ns)

ESL 260 L Eight-Cell Analysis		
Final Grade	Below Cut Score	At/Above Cut Score
A	0.00%	0.00%
B	0.00%	0.00%
C or CR	61.50%	74.50%
D,F,NC,WT	38.50%	25.50%
Totals	100.00%	100.00%

Correlation Between Success (1,2) and Cut Score Break (1,2) = (see 4-cell outcome)

Course is graded on a credit/no credit basis so results would be the same as the four-cell table.

ESL 260 L Partial Range Grades and all Scores	
A = 4	Score Range
B = 3	ESL 260L = 24 to 84
C or CR = 2	
D,F,NC,WT = 1	

Correlation Between Credit/No credit Grades (1,2,) and All 260L Test Scores = .149 (ns)

Caution: Sample size too small for definitive interpretation.

Summary Comment on ESL 260L: The sample size of 115 students is somewhat small. These preliminary results suggest that the Compass test is inadequate for use in this course. See Table 2 for summary.

Tables of Correlational Data

In Tables 1 and 2 are found all the basic correlations from the various courses plus additional correlations based on correction factors. It has been a belief by some individuals that if all W's (withdrew from course after census) were deleted from the analysis, the correlations would be acceptable. To check on this, the correlations were rerun after deleting all W's. Another suggestion about standard data analysis is that the correction for restriction of range should be applied, so that is included. Finally, it is well established that faculty grading variability is apt to alter any correlation value between scores and grades. In selected courses that correction was applied.

Table 1. Correlation Coefficients Between Compass Scores and Grades Under Varying Data Conditions For English, Math, Reading

Course	Original Correlation	Omitting All WTs	Correction Restricted Range	Standardized[2] Grades
English 1A				
4-cell =	-.038	-.009		
8-cell =	.114	.124		
Full r =	.027	.074	.053	.024
English 58				
4-cell =	.034	.064		
8-cell =	.113	.142		
Full r =	.111	.179	.184	.130
English 256				
4-cell =	.087	.050		
8-cell =	.097	.074		
Full r =	.099	.124	.201	.111
Math 51				
4-cell =	.166	.133		
8-cell =	.176	.154		
Mth1 Full r =	.173	.094		.211
Mth2 Full r =	.235	.157		.249
Math 53				
4-cell =	.138	.120		
8-cell =	.158	.150		
Mth2 Full r =	.198	.186		.244
Mth3 Full r =	.054	-.055		.041
Math 15				
4-cell =	.018	-.100		
8-cell =	-.114	-.169		
Mth4 Full r =	-.022	-.113		n/a
Mth5 Full r =	.065	-.015		n/a
English 4/5				
4-cell =	-.210	-.359		
8-cell =	-.228	-.376		
Full r =	-.015	-.140	-.023	-.065
English 71				
4-cell =	-.081	-.114		
8-cell =	-.145	-.184		
Full r =	-.133	-.136	-.305	-.130
English 268/270				
4-cell =	.112	.178		
8-cell =	.171	.235		
Full r =	.131	.157	.144	.082

[2]. Student letter grades were converted into z scores by the formula (X-Mean)/SD within each instructor. Thus an "A" earned from an instructor with a low grade-point average would result in a higher z value for a student than if earned from an instructor with a high grade-point average. The net effect is to give all instructors the same z average and standard deviation (mean = 0, SD = 1). This is one method to control for instructor grading variation, i.e., widely differing final grade standards.

Table 2. Correlation Coefficients Between Compass Scores and Grades For ESL

Course	Reading	Course	Writing	Course	Listening
ESL 2R		ESL 2W		ESL 2L	
4-cell =	-.078	4-cell =	.020	4-cell =	-.116
8-cell =	-.218	8-cell =	.116	8-cell =	-.173
Full r =	.370	Full r =	.098	Full r =	.153
ESL 1R		ESL 1W		ESL 1L	
4-cell =	-.255	4-cell =	-.174	4-cell =	.000
8-cell =	-.385	8-cell =	.453	8-cell =	.198
Full r =	-.207	Full r =	-.028	Full r =	.051
ESL 280R		ESL 280W		ESL 280L	
4-cell =	.124	4-cell =	-.008	4-cell =	.065
8-cell =	.276	8-cell =	.175	8-cell =	.098
Full r =	.277	Full r =	.118	Full r =	.277
ESL 270R		ESL 270W		ESL 270L	
4-cell =	n/a[3]	4-cell =	-.034	4-cell =	-.049
8-cell =	n/a	8-cell =	.159	8-cell =	.221
Full r =	.425	Full r =	.298	Full r =	.269
ESL 260R		ESL 260W		ESL 260L	
4-cell =	.347	4-cell =	.252	4-cell =	.093
8-cell =	n/a[4]	8-cell =	n/a	8-cell =	n/a
Full r =	.255	Full r =	.402	Full r =	.149

[3] No one scored below cut score.
[4] Course graded as credit/no-credit

BEST COPY AVAILABLE

Summary of the Compass Assessment Test Outcomes

As shown in Tables 1 and 2, the original correlations are either very low, non-existent, or even negative. The median value of all 27 full Pearson correlations (cut scores not considered) is .131 with a range of -.207 to .425. The unfortunate observation is that students taking the Compass test produce scores that have little or no relationship to final grades in these English writing, reading, mathematics, and ESL courses. If course entry decisions are based in part upon Compass test outcomes, then such decisions are of dubious value. Counselors who believe the Compass test should be given heavy weight are being misguided. Indeed, one could say that ARC is performing a disservice to students by practically insisting that all new students take the Compass test for use in academic counseling.

One may legitimately ask, why are there such low correlations? There are four likely possibilities:

1) What student efforts and skills go into producing a Compass test score are simply unrelated to the efforts and skills produced in the course. Final grades are based upon different things than what the test measures. Both may be good measures – they are just unrelated.

2) The Compass test is a poor test. A more refined instrument would be correlated with grades.

3) The Compass test is a good test. Instructors are grading on the wrong skills. Grades are not indicative of true learning or achievement.

4) The Compass test may be a good test, and instructors seem to be grading on achievement of the same type of skills. The problem is that one instructor will give an "A" for a specified mastery level while another instructor might grade the same level of mastery with a "C" – maybe even a "D." When you group such diverse grades for the same level of mastery, any correlation is apt to disappear. In short, final grades are inconsistent from one instructor to the next. With such grading inconsistency, no testing instrument can predict final grades. However, knowledge of instructors can be used to predict grades.

What is the correct answer? Probably "all of the above," meaning that there is some truth in each possibility. With instructor grading variation (#4) it is unlikely that all instructors of the same course are going to listen to a researcher complain about low correlations and therefore change their behavior. Some attempts have been made to change the whole concept of student assessment/placement because of this problem. Yet we do not believe that any new test is going to be much better than the Compass Test. So switching to another test is apt to be fruitless. What can be done?

One solution that is under discussion around the state is to have instructors of the same course write specific entrance objectives, that is, specific skills that all students must have as a prerequisite to enrolling in a course. Once those skills have been agreed upon and put in written form, the next task is to develop an accurate test that measures competency on those specific entrance objectives. Once normative data are collected on such a test, instructors can then decide on what will be the lowest score to gain admission to the course. With this type of placement decision, correlations can be run between the decision to enter or not enter and some type of course achievement. However, the focus of validity is not upon a correlation between scores and grades, but upon how well the test measures the course entrance standards.

Should the State Chancellor's Office remain in their present position of accepting only criterion validity or consequential validity, there may be a penalty (e.g. forfeiting matriculation dollars) because of implementing something different and hopefully, something better.

Implications for Planning

1. Immediately suspend assessment with the Compass test. At the very least, failing this recommendation, render all test prerequisites as "advisory only." The data do not support the use of the test scores as a prerequisite.

2. Request of the State Chancellor's Assessment Group the creation of criteria to allow for change or experimentation without monetary penalty.

3. Offer discussion sessions about this matter to all appropriate staff.

4. Have district wide committees develop entrance course standards.

5. Begin a search for a testing company or a select group of individuals who will produce a set of tests based upon the entrance standards.

6. Implement the new tests.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").